

Markov Random Processes are not Recoverable After Quantization and Mostly not Recoverable From Samples

Daniel Marco

Electrical Engineering Department
California Institute of Technology
Pasadena CA 91125, USA
Email: idaniel@ist.caltech.edu

Abstract—Markov random processes and general random processes are considered. It is shown that continuous-time, continuous-valued, wide-sense stationary, Markov random processes that have absolutely continuous second order distributions are not bandlimited. It is also shown that when these processes are strictly stationary and continuous almost surely, they cannot be recovered without error from their quantized versions. Further, it is shown that continuous-time, discrete-valued Markov random processes, which are uniformly bounded and satisfy an additional condition, can be recovered with zero average distortion from an appropriate set of samples for a general distortion measure. A similar result is shown for general continuous-time random processes with r^{th} power distortion measure. Additionally, it is shown that under a milder condition on the Markov processes and a different condition on the sampling times (e.g., uniform sampling), such processes cannot be recovered with zero average distortion.

Finally, the notion of information-singularity is extended to continuous-time random processes, and it is shown that both continuous- and discrete-time Markov processes are not information-singular.

I. INTRODUCTION

This paper considers fundamental questions mostly with regard to continuous-time Markov random processes that are either continuous- or discrete-valued, and considers also general continuous-time random processes. The results are concerned with the bandlimitedness, recoverability after quantization or from sets of samples, and information-singularity (or lack thereof for most categories) of such processes.

The first result of this paper, Theorem 1, is that continuous-time, continuous-valued, wide-sense stationary, Markov random processes with absolutely continuous second order distribution are not bandlimited. This result is important as it sheds light on the connection between two seemingly unrelated properties of random processes, Markovity and bandlimitedness, by showing that the former implies the latter.

The second result of this paper, Theorem 2, is motivated by the fact that in real life sample functions of random processes are not known with arbitrary precision. Instead, their values

are rounded via some kind of quantization that invariably introduces distortion, which in turn might render error-free recovery impossible. Nonetheless, error-free recovery might still be possible even in the presence of quantization. Specifically, when a continuous-time random process is passed through a quantizer, the quantized process can be used to deduce the precise locations of quantization threshold crossings of the original process. Some processes can be recovered without error from their threshold crossings [1]. Slepian [2] showed that such error-free recovery is not possible for an ergodic Gauss-Markov process and a binary quantizer with a threshold at the mean. Indeed, Theorem 2 shows that such error-free recovery is not possible for Markov processes in general.

The results that are concerned with the recoverability of random processes from sets of samples pertain to continuous-time, discrete-valued Markov processes and to other general random processes. Both positive and negative results are provided. Two positive results, Theorems 3 and 5, show that Markov processes that are uniformly bounded and satisfy an additional condition can be recovered with zero average distortion from an *appropriate* set of samples under a general distortion measure. Theorems 4 and 6 show that for the same conditions on the set of samples as in Theorems 3 and 5, respectively, general continuous-time random processes (i.e., processes that are not necessarily Markov and are either continuous- or discrete-valued) can be recovered with zero average distortion in the special case of r^{th} power distortion measure. The negative result, Theorem 7, shows that under a very mild condition on the Markov process, if there exists a gap between the sampling times that occurs with linear frequency (e.g., uniform sampling), then there exists a positive uniform lower bound to the distortion induced by any reconstruction scheme. Indeed, this is ordinarily the case.

Finally, the notion of information-singularity [3], [4] is extended to continuous-time random processes, and Theorem 9 shows that both continuous- and discrete-time Markov processes that are either continuous- or discrete-valued are not information-singular.

Some proofs are briefly sketched or omitted.

The remainder of this paper is organized as follows. Sec-

¹This work was supported by the Center for the Mathematics of Information at California Institute of Technology.

²This work was submitted to the IEEE transactions on Information Theory.

tion II introduces most of the notation and relevant background. Section III considers continuous-valued Markov processes and shows that these are neither bandlimited nor can be recovered after quantization. Section IV considers recoverability and non-recoverability from sets of samples of discrete-valued Markov processes and of general random processes. Section V shows that Markov processes are not information-singular. Lastly, Section VI offers concluding remarks.

II. BACKGROUND AND NOTATION

We represent a continuous-time random process X as the collection of real valued³ random variables $\{X_t(\omega), t \in (-\infty, \infty), \omega \in \Omega\}$ defined on the probability space (Ω, \mathcal{F}, P) . Let X_t denote the random variable at time t , and let $X(\omega)$ denote the sample path corresponding to ω .

The process X is said to be Markov if for any finite $n \geq 1$ and any $t_n < t_{n-1} < \dots < t_1 < s < t$, X_t is conditionally independent of $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ given X_s . In other words, the future is independent of the past given the present. This Markov property is sometimes called *unilateral* or *one-sided* Markov [5], [6].

We observe that if a Markov process has an absolutely continuous second order distribution, i.e., a second order probability density function (pdf), then it has an n^{th} order pdf for any finite $n \geq 1$.

A *bilateral* or *two-sided* Markov process is one where for any finite $m, n \geq 1$ and any $t_n < \dots < t_1 < t < \tau_1 < \dots < \tau_m$, X_t is conditionally independent of $X_{t_2}, \dots, X_{t_n}, X_{\tau_2}, \dots, X_{\tau_m}$ given X_{t_1} and X_{τ_1} . When second order pdfs exist, then, as mentioned, finite dimensional pdfs exist and it is easy to see using Bayes' Rule [5] that a unilateral Markov process is also a bilateral Markov process. If the Markov process is discrete-valued, then this follows directly from [5] without further conditions.

We define a quantizer to be a (measurable) function $Q : \mathbb{R} \rightarrow \mathbb{Z}$. The quantizer Q induces a partition of \mathbb{R} into cells $\{S_i\}$ such that $Q(x) = i$ when $x \in S_i$. Typically, these cells are intervals. A reconstruction of the process X at time t is a (measurable) function $g_t : A \rightarrow \mathbb{R}$, where A is some domain (we consider various reconstructions with different domains). Sometimes \hat{X} denotes a reconstruction of X . The operator $|\cdot|$ in addition to denoting absolute value denotes also the number of elements in a set or, if the set is an interval, it denotes its length. If the set contains several intervals, then it denotes the number of such intervals.

Next, we define two distortion measures. The first is $d(x, y) = \rho(|x - y|)$, where $\rho : [0, \infty) \rightarrow [0, \infty)$ is a nondecreasing (measurable) function for which $\rho(\alpha) > 0$ for $\alpha > 0$. The second is $\tilde{d} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, which is a nonnegative (measurable) function such that if $\alpha = \beta$, then $\tilde{d}(\alpha, \beta) = 0$, and if $B \subseteq \mathbb{R}$ is bounded, then there exists $S < \infty$ such that $\tilde{d}(\alpha, \beta) < S$ for all $\alpha, \beta \in B$.

We conclude this section with a brief discussion of discrete-valued Markov processes, which is relevant for Section IV and can be found in [7, pp. 231–233].

³Most of the results in this paper extend to more general metric spaces.

Let $S = \{1, 2, \dots, L\}$, $1 \leq L \leq \infty$ be the state space of a discrete-time, discrete-valued Markov chain Y . Let $V = \{v_i\}_{i=1}^L$ be a set of real numbers such that $v_i \neq v_j$ if $i \neq j$. The random process X is a continuous-time, discrete-valued, Markov process if $X_t \in V$ for all t , where we say that X_t is in state i when $X_t = v_i$, such that the amount of time X spends in state i before transitioning to another state is exponentially distributed with rate λ_i . The state to which the process transitions is determined according to the Markov chain Y , which is called *the underlying* Markov chain. The exponential parameters range from 0 to ∞ . If $\lambda_i = 0$, we say that state i is *absorbing*, namely, once X enters state i , it never leaves it, i.e., if $X_t = v_i$, then $X_s = v_i$ for all $s > t$ with probability one. If $\lambda_i = \infty$, then state i is called *instantaneous*, since once entered it is instantaneously left. If X is stationary, then we let $q_i \triangleq \Pr(X_t = v_i)$, which is ordinarily different from the probability that Y is in state i .

The state space S of the Markov chain Y can be partitioned into sets called *communicating classes* such that any two states in the same communicating class have a positive probability to transition from one to the other in a finite number of transitions [7, pp. 167–174].

III. CONTINUOUS-VALUED MARKOV PROCESSES

The following theorem shows that continuous-valued Markov processes are not bandlimited.

Theorem 1: A continuous-time, wide-sense stationary, Markov random process with absolutely continuous second order distribution and finite second moment is not bandlimited.

Proof: If a wide-sense stationary random process X has finite second moment and is bandlimited, then the sampling theorem [8] shows that there exists a sampling interval $\tau > 0$ such that given samples at times $n\tau$, $n \in \mathbb{Z}$, the process can be reconstructed in the mean-square sense for any time. Namely, for any time t

$$\lim_{N \rightarrow \infty} E(X_t - \tilde{X}_{N,t})^2 = 0,$$

where $\tilde{X}_{N,t}$ is an appropriately chosen function of t and the samples of X at times $\{-N\tau, \dots, N\tau\}$. To show that the given Markov process, denoted by X , is not bandlimited we consider the contrapositive. Namely, we show that for any sampling interval $\tau > 0$ and any reconstruction procedure there exists a time s for which mean-square convergence fails. This implies that X is not bandlimited.

Let $\tau > 0$ be given. Set $\hat{X}_{N,t} = E[X_t | X_{-N\tau}, \dots, X_{N\tau}]$ and observe that it is the best mean-square estimate of X_t given the samples $\{X_{-N\tau}, \dots, X_{N\tau}\}$. Set $s = \frac{\tau}{2}$ and obtain

$$\liminf_{N \rightarrow \infty} E(X_s - \hat{X}_{N,s})^2 = E(X_s - E(X_s | X_0, X_\tau))^2, \quad (1)$$

where the equality is due to the fact that X is a bilateral Markov process since it has a second order pdf.

Finally, the right-hand side of (1) is positive since X_0, X_s , and X_τ have an absolutely continuous joint distribution, which implies that $\Pr(X_s = E(X_s | X_0, X_\tau)) < 1$. \square

The next theorem shows that when the Markov process above is stationary and continuous almost surely, it cannot be reconstructed without error from its quantized version. We comment that stationary and separable Gauss-Markov processes are an example of processes that are continuous almost surely (in fact they are almost surely sample path continuous) and have absolutely continuous second order distribution.

Theorem 2: Let X be a continuous-time, stationary Markov random process that is continuous almost surely and has absolutely continuous second order distribution. Let Q be an arbitrary quantizer that has a cell that contains an interval and the interval occurs with positive probability. Let X_Q denote the quantized version of X , i.e., $X_{Q,t} = Q(X_t)$. There exists $\varepsilon > 0$ such any stationary reconstruction of X from X_Q incurs distortion (with respect to $d(\cdot, \cdot)$) greater than ε . Specifically,

$$D(X, g(X_Q)) \triangleq \liminf_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[d(X_t, g_t(X_Q))] dt > \varepsilon ,$$

where g is any stationary reconstruction of X from $\{X_{Q,t}\}_{t=-\infty}^{\infty}$ (i.e., g has as input the whole quantized process) and $g_t(X_Q)$ is its value at time t .

Proof Sketch: Let $W = (a, b)$ be an interval that occurs with positive probability and is contained by a quantization cell of Q . It can be shown, using the facts that X has an absolutely continuous distribution and is continuous almost surely, that there exist $\tau > 0$ and a corresponding set $G_\tau \subseteq \Omega$ such that $P(G_\tau) > 0$ and $X_t(\omega) \in W$ for all $0 \leq t \leq \tau$ and all $\omega \in G_\tau$.

Set $s = \frac{\tau}{2}$. There exist reconstructions of X at time t , $\hat{g}_t(X_Q, X_0, X_\tau)$, $\hat{g}_t(\{X_Q\}_{t=0}^\tau, X_0, X_\tau)$, and $\tilde{g}_t(X_0, X_\tau)$ that use $\{X_Q, X_0, X_\tau\}$, $\{\{X_{Q,t}\}_{t=0}^\tau, X_0, X_\tau\}$, and $\{X_0, X_\tau\}$, respectively, such that

$$\begin{aligned} D(X, g(X_Q)) &\stackrel{(a)}{=} E[d(X_s, g_s(X_Q))] \\ &\stackrel{(b)}{\geq} E[d(X_s, \tilde{g}_s(X_Q, X_0, X_\tau))] \\ &\stackrel{(c)}{=} E[d(X_s, \hat{g}_s(\{X_Q\}_{t=0}^\tau, X_0, X_\tau))] \\ &\stackrel{(d)}{\geq} E[d(X_s, \tilde{g}_s(X_0, X_\tau)) | G_\tau] P(G_\tau) , \end{aligned}$$

where (a) follows from the stationarity of X and g , (b) is due to \tilde{g} having a larger parameter input set, (c) derives from X being a bilateral Markov process⁴ since it has a second order pdf, and (d) follows since $\{X_{Q,t}\}_{t=0}^\tau$ is constant given G_τ .

Next, it can be shown that since the distortion function ρ (which defines $d(\cdot, \cdot)$) is nondecreasing, there exists a bounded distortion function $\bar{\rho} : [0, \infty) \rightarrow [0, M]$, $M < \infty$ that is continuous and nondecreasing such that $\bar{\rho}(\alpha) \leq \rho(\alpha)$ for $\alpha \in [0, \infty)$ and $\bar{\rho}(\alpha) > 0$ for $\alpha > 0$. Thus it suffices to show that

$$\begin{aligned} \inf_{\tilde{g}_s} E[\rho(|X_s - \tilde{g}_s(X_0, X_\tau)|) | G_\tau] P(G_\tau) \\ \geq \inf_{\tilde{g}_s} E[\bar{\rho}(|X_s - \tilde{g}_s(X_0, X_\tau)|) | G_\tau] P(G_\tau) > 0 . \end{aligned}$$

⁴Given X_0 and X_τ , X_s is independent of the uncountable collection $\{X_t\}_{t < 0, t > \tau}$. While typically such independence is stated relative to a countable collection, it holds in the uncountable case too [9, pp. 393–395].

The second inequality is seen as follows. Since $\bar{\rho}$ is a non-decreasing function, it can be shown that for any observed values of X_0 and X_τ , the infimum can be taken over a closed and bounded interval, i.e., over a compact set. Since $\bar{\rho}$ is continuous and bounded, the expectation $E[\bar{\rho}(|X_s - \tilde{g}_s(X_0, X_\tau)|) | G_\tau]$ is continuous, and therefore the infimum is a minimum. Let \tilde{g}_s^* be an estimator that attains the second infimum in last equation. All that remains to show is that

$$E[\bar{\rho}(|X_s - \tilde{g}_s^*(X_0, X_\tau)|) | G_\tau] P(G_\tau) > 0 ,$$

which holds as similarly shown in the proof of Theorem 1. \square

IV. RECOVERABILITY FROM SAMPLES

This section shows that discrete-valued Markov random processes with distortion measure $\tilde{d}(\cdot, \cdot)$ and general random processes with r^{th} power distortion measure can be recovered with zero average distortion from appropriate sets of samples. It also shows that Markov processes cannot be recovered with zero average under ordinary conditions on the set of samples.

Recalling the definitions of λ_i and q_i we have the following theorem.

Theorem 3: Let X be a continuous-time, discrete-valued (with $1 \leq L \leq \infty$ states), stationary, separable, Markov random process states that is uniformly bounded so that $\tilde{d}(X_{t_1}, X_{t_2}) \leq S$ for all t_1, t_2 and some $S < \infty$. Let $M < \infty$ be such that $\sum_{i=1}^L q_i \lambda_i \leq M$. If $\{t_k\}_{k=-\infty}^{\infty}$ is a set of sampling times for which $t_{k+1} > t_k$, $\lim_{k \rightarrow \infty} t_k = \infty$, $\lim_{k \rightarrow -\infty} t_k = -\infty$, and $\lim_{k \rightarrow \pm\infty} (t_{k+1} - t_k) = 0$, then there exists a reconstruction \hat{X} of X from the samples at times $\{t_k\}_{k=-\infty}^{\infty}$ such that

$$D(X, \hat{X}) \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[\tilde{d}(X_t, \hat{X}_t)] dt = 0 . \quad (2)$$

Proof Sketch: The main idea is to choose $\hat{X}_t = X_{t_k}$ whenever $t \in [t_k, t_{k+1})$, namely, “sample and hold”, and show that this choice results in zero average distortion. We first bound the probability that \hat{X}_t does not equal X_t . Specifically, let $\varepsilon_k \triangleq t_{k+1} - t_k$ and consider $t \in [t_k, t_{k+1})$. It can be shown, using the fact that X is Markov and thus spends an exponential amount of time in any state, that

$$\Pr(\hat{X}_t \neq X_t) < \sum_{i=1}^L q_i \lambda_i \varepsilon_k \leq M \varepsilon_k .$$

Recalling that $\tilde{d}(X_{t_1}, X_{t_2}) \leq S$ for all t_1 and t_2 , it follows that $E[\tilde{d}(X_t, \hat{X}_t)] \leq S M \varepsilon_k$ for $t \in [t_k, t_{k+1})$.

Next fix k , let $n > 0$ be such that $T \in (t_{k+n}, t_{k+n+1}]$, and set $\bar{\varepsilon}_k = \max\{\varepsilon_k, \varepsilon_{k+1}, \dots\}$. It can be shown that

$$\int_{t_k}^T E[\tilde{d}(X_t, \hat{X}_t)] dt \leq S M \sum_{i=0}^n \varepsilon_{k+i}^2 \leq S M \bar{\varepsilon}_k (T - t_k + \bar{\varepsilon}_k) .$$

Similarly, fix m , let $l > 0$ be such that $-T \in [t_{m-l}, t_{m-l+1})$, and set $\bar{\varepsilon}_m \triangleq \max\{\varepsilon_m, \varepsilon_{m-1}, \dots\}$. It can be shown that

$$\int_{-T}^{t_m} E[\tilde{d}(X_t, \hat{X}_t)] dt \leq S M \bar{\varepsilon}_m (T + t_m + \bar{\varepsilon}_m) .$$

Finally, using the last two equations

$$\begin{aligned}
D(X, \hat{X}) &= \lim_{m \rightarrow -\infty} \lim_{k \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{1}{2T} \left[\int_{t_m}^{t_k} E[\tilde{d}(X_t, \hat{X}_t)] dt \right. \\
&\quad \left. + \int_{-T}^{t_m} E[\tilde{d}(X_t, \hat{X}_t)] dt + \int_{t_k}^T E[\tilde{d}(X_t, \hat{X}_t)] dt \right] \\
&\leq \lim_{m \rightarrow -\infty} \lim_{k \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{1}{2T} \left[S(t_k - t_m) \right. \\
&\quad \left. + SM\tilde{\varepsilon}_k(T - t_k + \tilde{\varepsilon}_k) + SM\tilde{\varepsilon}_m(T + t_m + \tilde{\varepsilon}_m) \right] \\
&= \frac{SM}{2} \lim_{m \rightarrow -\infty} \lim_{k \rightarrow \infty} (\tilde{\varepsilon}_k + \tilde{\varepsilon}_m) = 0.
\end{aligned}$$

□

Theorem 4: Let X be a continuous-time, wide-sense stationary random process that is uniformly continuous in r^{th} mean and has finite r^{th} moment if $1 \leq r < \infty$ and finite mean if $0 < r < 1$. If $\{t_k\}_{k=-\infty}^{\infty}$ is a set of sampling times for which $t_{k+1} > t_k$, $\lim_{k \rightarrow \infty} t_k = \infty$, $\lim_{k \rightarrow -\infty} t_k = -\infty$, and such that for some $0 < \alpha < \beta < \infty$, $\lim_{k \rightarrow \pm\infty} (t_{k+1} - t_k) = 0$, then there exists a reconstruction \hat{X} of X from the samples at times $\{t_k\}_{k=-\infty}^{\infty}$ such that

$$D(X, \hat{X}) \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E|X_t - \hat{X}_t|^r dt = 0.$$

Proof: We choose $\hat{X}_t = X_{t_k}$ whenever $t \in [t_k, t_{k+1})$, and show that this choice results in zero average distortion.

Let $\varepsilon > 0$ be given. Since X is continuous in r^{th} mean there exists τ_ε such that for all $\tau < \tau_\varepsilon$, $E|X_t - X_{t+\tau}|^r < \varepsilon$. This, the choice of \hat{X} , and the fact that $\lim_{k \rightarrow \pm\infty} (t_{k+1} - t_k) = 0$ imply that there exist K_ε^+ and K_ε^- such that $E|X_t - \hat{X}_t|^r < \varepsilon$ for all $t > t_{K_\varepsilon^+}$ and $t < t_{K_\varepsilon^-}$. Letting $\sigma^r \triangleq E|X_t|^r < \infty$ and using $E|X_t| < \infty$ if $0 < r < 1$, we obtain

$$\begin{aligned}
D(X, \hat{X}) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \left[\int_{K_\varepsilon^-}^{K_\varepsilon^+} E|X_t - \hat{X}_t|^r dt \right. \\
&\quad \left. + \int_{-T}^{K_\varepsilon^-} E|X_t - \hat{X}_t|^r dt + \int_{K_\varepsilon^+}^T E|X_t - \hat{X}_t|^r dt \right] \\
&\leq \lim_{T \rightarrow \infty} \frac{1}{2T} \left[(K_\varepsilon^+ - K_\varepsilon^-) 2^r \max\{(E|X_t|)^r, \sigma^r\} \right. \\
&\quad \left. + (K_\varepsilon^- + T)\varepsilon + (T - K_\varepsilon^+)\varepsilon \right] = \varepsilon,
\end{aligned}$$

where the inequality follows since $E|X_t - \hat{X}_t|^r \leq E(|X_t| + |\hat{X}_t|)^r$ and since $(a+b)^r \leq 2^{r-1}(a^r + b^r)$ for $a, b \geq 0$, $r \geq 1$, which is obtained using the convexity of x^r for $r \geq 1$, and from the fact that x^r is concave for $0 < r < 1$, in which case $E(|X_t| + |\hat{X}_t|)^r \leq (E|X_t| + E|\hat{X}_t|)^r = 2^r(E|X_t|)^r$. Since ε is arbitrary the result follows. □

The condition on the sampling times in Theorems 3 and 4 can be relaxed somewhat so as to still attain zero average distortion. Specifically, as the next two theorems show, such is the case when allowing bounded (above and below) gaps between sampling times, whose frequency is sub-linear. Before proceeding to the next theorems we define the $\lim^{\alpha, \beta}$ notation.

Definition 1: Let $\{a_k\}_{k=-\infty}^{\infty}$ be a sequence of nonnegative real numbers. Let $0 < \alpha < \beta < \infty$. Let $I_N = \{-|N| \leq$

$k \leq |N| : \alpha \leq a_k \leq \beta\}$. If $\lim_{k \rightarrow \infty} a_k = v$ and $\lim_{N \rightarrow \infty} \frac{|I_N|}{N} = 0$, then we say that $\lim_{k \rightarrow \infty}^{\alpha, \beta} a_k = v$. $\lim_{k \rightarrow -\infty}^{\alpha, \beta} a_k$ is analogously defined.

Theorem 5: Let X be a random process as in Theorem 3 that satisfies the same conditions. If $\{t_k\}_{k=-\infty}^{\infty}$ is a set of sampling times for which $t_{k+1} > t_k$, $\lim_{k \rightarrow \infty} t_k = \infty$, $\lim_{k \rightarrow -\infty} t_k = -\infty$, and such that for some $0 < \alpha < \beta < \infty$, $\lim_{k \rightarrow \pm\infty}^{\alpha, \beta} (t_{k+1} - t_k) = 0$, then there exists a reconstruction \hat{X} of X from the samples at times $\{t_k\}_{k=-\infty}^{\infty}$ such that

$$D(X, \hat{X}) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[\tilde{d}(X_t, \hat{X}_t)] dt = 0.$$

Theorem 6: Let X be a continuous-time, wide-sense stationary random process that is uniformly continuous in r^{th} mean and has finite r^{th} moment if $1 \leq r < \infty$ and finite mean if $0 < r < 1$. If $\{t_k\}_{k=-\infty}^{\infty}$ is a set of sampling times for which $t_{k+1} > t_k$, $\lim_{k \rightarrow \infty} t_k = \infty$, $\lim_{k \rightarrow -\infty} t_k = -\infty$, and such that for some $0 < \alpha < \beta < \infty$, $\lim_{k \rightarrow \pm\infty}^{\alpha, \beta} (t_{k+1} - t_k) = 0$, then there exists a reconstruction \hat{X} of X from the samples at times $\{t_k\}_{k=-\infty}^{\infty}$ such that

$$D(X, \hat{X}) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E|X_t - \hat{X}_t|^r dt = 0.$$

We conclude this section with a theorem that shows that if the sampling times of a discrete-valued Markov process have a gap $\delta > 0$ that occurs with linear frequency (e.g., uniform sampling), then there exists a positive uniform lower bound to the distortion induced by any reconstruction scheme.

Theorem 7: Let X be a continuous-time, discrete-valued, stationary, separable, Markov random process whose underlying Markov chain has a communicating class, which has at least two states and no absorbing or instantaneous states, that occurs with positive probability. Let $\{t_k\}_{k=-\infty}^{\infty}$ be such that $t_{k+1} > t_k$, $\lim_{k \rightarrow \infty} t_k = \infty$, and $\lim_{k \rightarrow -\infty} t_k = -\infty$, and let $l_k = (t_k, t_{k+1})$. For $\delta > 0$ let $I_T^\delta = \{l_k : |l_k| \geq \delta, t_{k+1} \leq t_k, t_k \geq -T\}$. If there exists $\delta > 0$ such that $\liminf_{T \rightarrow \infty} \frac{|I_T^\delta|}{T} > 0$, then there exists $\varepsilon > 0$ such that any reconstruction of X from the samples at times $\{t_k\}_{k=-\infty}^{\infty}$ incurs distortion (with respect to $d(\cdot, \cdot)$) greater than ε . Specifically,

$$\begin{aligned}
D(X, g(\{X_{t_k}\}_{k=-\infty}^{\infty})) \\
\triangleq \liminf_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[d(X_t, g_t(\{X_{t_k}\}_{k=-\infty}^{\infty}))] dt > \varepsilon,
\end{aligned}$$

where g is any reconstruction of X from the samples at times $\{t_k\}_{k=-\infty}^{\infty}$ and $g_t(\cdot)$ is its value at time t .

V. INFORMATION-SINGULARITY

Information-singularity was first introduced by Berger [3], who coined discrete-time random processes that generate information at a sub-linear rate as information-singular. More precisely, a process is *block (variable-length) information-singular* if there exists a sequence of block (variable-length) codes $C_n = (m = 2n + 1, e_m, d_m)$, where e_m is an encoding function and d_m is a decoding function, such that when C_n is applied to $\mathbf{X}_m \triangleq (X_{-n}, \dots, X_n)$, the rate (average rate)

and average per-letter distortion converge to zero as n tends to infinity. The expression for the convergence of the per-letter distortion is given by

$$\lim_{n \rightarrow \infty} \frac{E \sum_{k=-n}^n (X_k - \hat{X}_k)^2}{2n+1} = 0,$$

where $\hat{X}_m = d_m(e_m(\mathbf{X}_m))$.

We extend information-singularity to continuous-time random processes. Let $\{X_t\}$ denote a continuous-time random process, let $m = 2n$, let C^m be the range of possible values for $\mathbf{X}_m \triangleq \{X_t\}_{t=-n}^n$, let $\{0, 1\}^*$ be the set of all finite binary strings, and let $l(s)$ denote the length of $s \in \{0, 1\}^*$.

Definition 2: $\{X_t\}$ is *block information-singular* if for all $R > 0$, there exist $\{e_m : C^m \rightarrow \{1, 2, \dots, [2^{mR}]\}\}$ and $\{d_m : \{1, 2, \dots, [2^{mR}]\} \rightarrow C^m\}$ such that

$$\lim_{n \rightarrow \infty} \frac{E \int_{-n}^n (X_t - \hat{X}_t)^2 dt}{2n} = 0, \quad (3)$$

where $\hat{X}_m = d_m(e_m(\mathbf{X}_m))$.

Definition 3: $\{X_t\}$ is *variable-length information-singular* if for all $R > 0$, there exist $\{e_m : C^m \rightarrow \{0, 1\}^*\}$ and $\{d_m : \{0, 1\}^* \rightarrow C^m\}$ such that $\lim_{m \rightarrow \infty} m^{-1} E[l(e_m(\mathbf{X}_m))] < R$ and (3) holds when $\hat{X}_m = d_m(e_m(\mathbf{X}_m))$.

We say that a process is not information-singular if it is neither block nor variable-length information-singular.

We note that Berger [3] introduced information-singularity with respect to mean-squared error over the space \mathbb{R} , and Hajek [4] considered information-singularity in more generality for an arbitrary complete separable metric space (B, d) with metric d and with respect to distortion measure $d^\alpha(a, b)$, where $\alpha > 0$ and $a, b \in B$. To keep notation simple, we restrict discussion to mean-squared error over \mathbb{R} .

We quote below Theorem 2 from [4], which will be used to show the main result of this section.

Theorem 8: [4] If X is stationary (discrete-time) random process with a complete separable state space (B, d) and if there exists $b^* \in B$ such that, for all k , $E[d^\alpha(X_k, b^*)] < \infty$, which we refer to as the *Integrability Condition*, then the following are equivalent:

- a) X is block information-singular
- b) X is variable-length information-singular
- c) X has zero entropy rate

Theorem 9: If X is a stationary, measurable Markov random process with finite second moment that is not constant with probability one, then X is not information-singular.

Proof: If X is a discrete-time random process, then since it is not constant with probability one, $H(X) = H(X_2|X_1) > 0$, where $H(X)$ is the entropy rate of X and X_n is the value of X at time n . Thus Theorem 8, which can be used since $EX^2 < \infty$, implies that X is not information-singular.

Next, suppose X is a continuous-time random process. Consider the discrete-time random process Y defined by $Y_k = \{X_t\}_{t=k}^{k+1}$. Since X is measurable and has finite second moment, its sample functions are square integrable over any finite interval with probability one. Thus Y is a discrete-time

random process over the complete separable metric space L_2 of functions over the interval $[0, 1]$. Since X is Markov, so is Y . Further, since X has finite second moment, Y satisfies the integrability condition of Theorem 8, since $E \int_k^{k+1} (X_t - 0)^2 dt = EX_0^2 < \infty$. Therefore, the first part of the theorem implies that Y is not information-singular.

Next, we show by contradiction that X is not information-singular. Specifically, we assume that X is (block or variable-length) information-singular and show that this implies that Y is information-singular, thus arriving at a contradiction.

Let $\{e_m, d_m\}$, $m = 2n$ be a sequence of encoders and decoders that satisfy any rate constraint such that

$$\lim_{n \rightarrow \infty} \frac{E \int_{-n}^n (X_t - \hat{X}_t)^2 dt}{2n} = 0.$$

This implies that $\lim_{n \rightarrow \infty} \frac{E \int_{-n}^{n+1} (X_t - \hat{X}_t)^2 dt}{2n+1} = 0$. Letting the encoders and decoders for the discrete-time process Y be the same as those for the continuous-time process X and denoting $\hat{Y}_k = \{\hat{X}_t\}_{t=k}^{k+1}$, it follows that

$$\frac{E \sum_{k=-n}^n (Y_k - \hat{Y}_k)^2}{2n+1} = \frac{E \int_{-n}^{n+1} (X_t - \hat{X}_t)^2 dt}{2n+1} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

This implies that Y is information-singular as desired. \square

VI. CONCLUSIONS

It was shown that continuous-time, continuous-valued Markov random processes are not bandlimited and that they cannot be recovered without error after quantization. It was shown further that under certain conditions on the sets of sampling times, continuous-time, discrete-valued Markov processes can be recovered with zero average distortion with respect to the general distortion measure $\tilde{d}(\cdot, \cdot)$. This was also shown for general continuous-time random processes with r^{th} power distortion measure. Additionally, it was shown that under more ordinary conditions on the set of sampling times (e.g., uniform sampling) continuous-time, discrete-valued Markov processes cannot be recovered with zero average distortion. Finally, information-singularity was extended to continuous-time random processes and it was shown that Markov processes are not information-singular.

REFERENCES

- [1] I. Bar-David, "An implicit sampling theorem for bounded bandlimited functions," *Info. Contr.*, vol. 24, pp. 36–44, Jan. 1974.
- [2] D. Slepian, "Estimation of the Gauss-Markov process from observation of its sign," *Stoch. Proc. Appl.*, vol. 14, pp. 249–265, 1983.
- [3] T. Berger, "Information-Singular random processes," *IEEE Trn. Inf. Thry.*, vol. 21, pp. 502–511, Sep. 1975.
- [4] B. Hajek, "Information-Singularity and recoverability of random processes," *IEEE Trn. Inf. Thry.*, vol. 28, pp. 422–429, May 1982.
- [5] K. Abend, T. J. Harley and L. N. Kanal, "Classification of binary random patterns," *IEEE Trn. Inf. Thry.*, vol. 11, pp. 538–544, Oct. 1965.
- [6] H. Derin and P. A. Kelly, "Discrete-Index Markov-type random processes," *Proc. of IEEE*, vol. 77, pp. 1485–1509, Oct. 1989.
- [7] S. Ross, *Stochastic Processes*, 2nd Ed. New York, Wiley & Sons, 1981.
- [8] A. V. Balakrishnan, "A note on the Shannon sampling principle for continuous signals," *IRE Trn. Inf. Thry.*, vol. 3, pp. 143–146, June 1957.
- [9] H. Bauer, *Probability Theory and Elements of Measure Theory*, 2nd Ed. New York, Academic Press, 1981.